

TỔNG HỢP VÀ BIỂU HIỆN PROTEIN CAS9 TÁI TỔ HỢP TRONG TẾ BÀO VI KHUẨN *ESCHERICHIA COLI*

Hoàng Xuân Cường¹, Đỗ Như Bình²

TÓM TẮT

Mục tiêu: Tổng hợp nhân tạo gen mã hóa Cas9 bằng công nghệ Golden gate và biểu hiện protein tái tổ hợp trong tế bào vi khuẩn *E. coli*. **Đối tượng và phương pháp:** Nghiên cứu thực nghiệm, gen mã hóa protein Cas9 được tổng hợp bằng phương pháp Golden gate, biến nạp vào vector pET52b và biểu hiện ở tế bào *E. coli*. Tinh sạch protein bằng sắc ký ái lực Ni-NTA. **Kết quả:** Tạo thành công dòng tế bào *E. coli* BL21(DE3) mang vector pET52b-Cas9 có khả năng biểu hiện protein Cas9 tái tổ hợp ở pha tan. Thu nhận được protein Cas9 tái tổ hợp với nồng độ 4,1 mg/ml và độ tinh sạch kiểm tra bằng SDS-PAGE đạt $\geq 98\%$. **Kết luận:** Đã tạo được nguồn cung cấp protein Cas9 tái tổ hợp phục vụ cho các nghiên cứu trong lĩnh vực chỉnh sửa gen.

* Từ khóa: Protein Cas9 tái tổ hợp; Chỉnh sửa gen; Biểu hiện.

Synthesis and Expression of Recombinant Cas9 Protein in Escherichia Coli

Summary

Objectives: To artificially synthesize Cas9-encoding gene by Golden gate technology and to express a recombinant protein in *E. coli* bacteria cells. **Subjects and methods:** An experimental study, the gene encoding Cas9 protein was synthesized by Golden gate method, transformed into pET52b vector, and expressed in *E. coli* cells. Purify proteins by Ni-NTA affinity chromatography. **Results:** Successfully constructed *E. coli* BL21(DE3) cell line carrying pET52b-Cas9 vector with a capable of expressing recombinant Cas9 protein insoluble phase. Recombinant Cas9 protein was obtained at a concentration of 4.1 mg/mL, and the purity was above 98% that verified by SDS-PAGE. **Conclusion:** Self-production and provided an abundant source of recombinant Cas9 protein for research in the field of gene editing.

* Keywords: Recombinant Cas9 protein; Gene editing; Expression.

¹Học viện Quân y

²Ban Khoa học Quân sự, Bệnh viện Quân y 103, Học viện Quân y

Người phản hồi: Hoàng Xuân Cường (hoangxuancuong@vmmu.edu.vn)

Ngày nhận bài: 05/5/2021

Ngày bài báo được đăng: 26/5/2021

ĐẶT VẤN ĐỀ

Trong suốt thập kỷ qua, các kỹ thuật thao tác DNA bộ gen dựa trên hoạt động của enzyme endonuclease được sử dụng rộng rãi như Zinc Finger Nucleases (ZFNs) và Transcription Activator Like Effector Nuclease (TALEN) [1]. Tuy nhiên, các phương pháp này có quy trình thiết kế và việc sử dụng trong thực tế rất phức tạp, do đó tính ứng dụng không cao, đặc biệt trong lâm sàng. Hiện nay, nhiều nghiên cứu đã sử dụng thành công kỹ thuật biến đổi DNA bộ gen dựa trên hệ thống CRISPR/Cas [2, 3, 4]. Hệ thống này dựa trên cơ chế “miễn dịch” của vi khuẩn chống lại sự xâm nhiễm phân tử DNA ngoại lai từ virus hoặc DNA plasmid [6, 7]. Khác với hệ thống “miễn dịch” dựa trên enzyme cắt giới hạn, hệ thống này dựa trên phân tử RNA để nhận diện và phá hủy DNA ngoại lai. Để bảo vệ vi khuẩn khỏi DNA ngoại lai, vi khuẩn cần được gắn chèn một đoạn ngắn DNA ngoại lai vào DNA bộ gen tại vùng trình tự lặp lại (CRISPR - Clustered Regularly Interspaced Short Palindromic Repeat) [5]. Vùng trình tự này được phiên mã và xử lý thành các đoạn RNA ngắn (gọi là crRNA-CRISPR-derived RNA), các crRNA này liên kết với endonuclease Cas để nhận diện DNA mục tiêu và cắt phân tử DNA mục tiêu thông qua hoạt động endonuclease của protein Cas [1, 5].

Gần đây, một công cụ mới dựa trên Cas9 (Protein 9 có hoạt tính nuclease liên quan đến CRISPR của vi khuẩn) có nguồn gốc từ *Streptococcus pyogenes*

SF370 được ứng dụng nhiều để điều chỉnh chức năng gen, sửa chữa những sai sót trong bộ gen của vi sinh vật [5, 7], như làm đảo đoạn hoặc chuyển vị trí, tác động đến vùng protein nhằm điều hòa quá trình phiên mã, biến đổi di truyền ngoại gen và hiển thị các locut gen mong muốn... Trong nghiên cứu này, chúng tôi tiến hành tổng hợp nhân tạo gen mã hóa protein Cas9, biểu hiện trong tế bào vi khuẩn *E. coli* và tinh sạch protein Cas9 tái tổ hợp, định hướng sử dụng trong thiết lập hệ thống CRISPR/Cas9 ứng dụng trong lĩnh vực chỉnh sửa gen điều trị bệnh lý di truyền.

ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

1. Đối tượng nghiên cứu

- Gen mã hóa protein Cas9.

* *Vật liệu, hóa chất và thiết bị nghiên cứu:*

- Vật liệu: Chủng *E. coli* Top10 được dùng để tách dòng gen mã hóa cho Cas9; hệ vector pJET1.2, pET52b (hãng Thermo Fisher Scientific, Mỹ) dùng để tách dòng gen; các cặp mồi và các đoạn oligo để tổng hợp các mảnh gen mã hóa cho Cas9; cặp mồi pJET1.2-Forward và pJET1.2-Reverse dùng để sàng lọc chọn dòng và giải trình tự.

- Hóa chất: Các kit tinh sạch trực tiếp sản phẩm PCR (hãng Thermo Fisher Scientific, Mỹ), kit tinh sạch trên gel sản phẩm PCR (hãng NEB, Mỹ); hóa chất điện di EDTA High Ranger 1kb DNA

Ladder (hãng Norgen Biotek Corp, Canada), ExceLBand 50bp DNA Ladder; enzyme T4 DNA ligase; hóa chất Western blot; hóa chất điện di protein SDS-PAGE.

- Thiết bị nghiên cứu: Máy chu kỳ nhiệt 96 mẫu (hãng Eppendorf, Đức); máy soi chụp ảnh gel (hãng Gensnap, Mỹ); bộ điện di DNA Horizontal mini (hãng CBS Scientific, Mỹ); máy quang phổ NanoDrop 2000 (hãng Thermo Fisher Scientific, Mỹ); máy lắc ổn nhiệt DTU-2C (hãng Taitec, Nhật Bản); tủ lạnh sâu -20°C và -80°C (hãng Sanyo, Nhật Bản)...

- Địa điểm nghiên cứu: Phòng Công nghệ Gen, Viện Nghiên cứu Y Dược học Quân sự, Học viện Quân y.

- Thời gian nghiên cứu: 8/2019 - 9/2020.

2. Phương pháp nghiên cứu

* *Thiết kế nghiên cứu*: Nghiên cứu thực nghiệm.

* *Kỹ thuật sử dụng*:

- Tổng hợp gen mã hóa Cas9 bằng công nghệ Golden gate:

+ Tối ưu trình tự gen: Dựa trên trình tự gen Cas9 trên GenBank, sử dụng công cụ tìm kiếm chuỗi bắt cặp (BLAST) và công cụ phân tích mã bộ ba hiếm (GenScript) để tối ưu bộ ba mã hóa cho chuỗi amino acid.

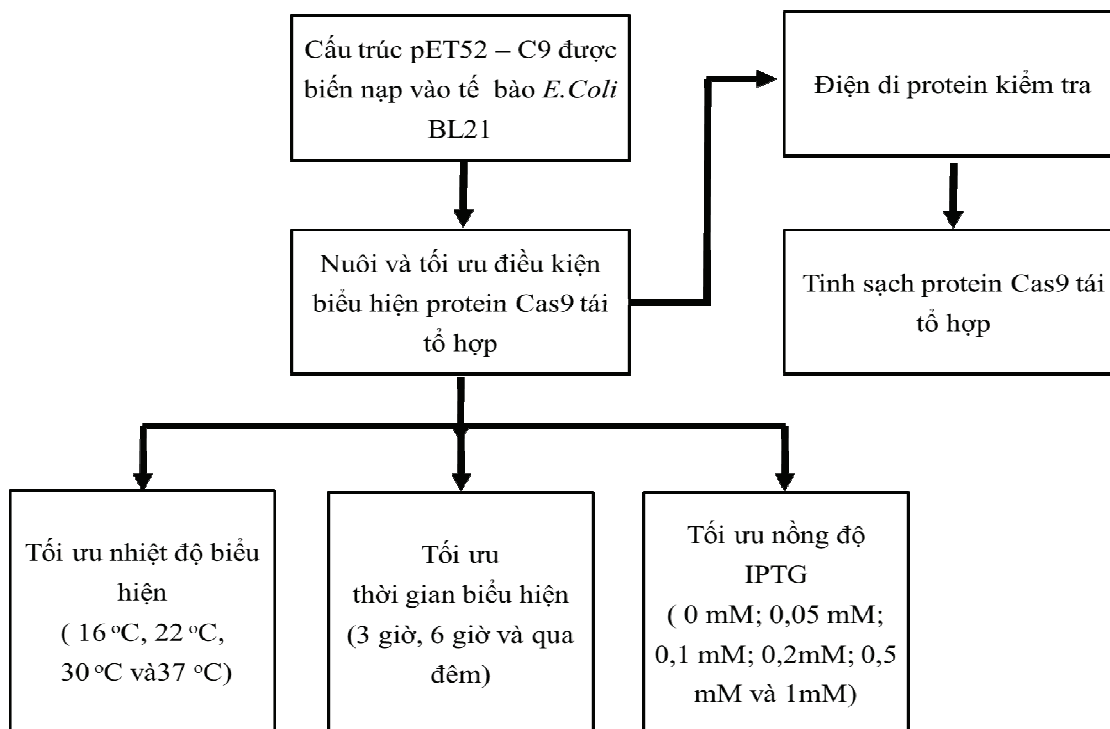
+ Thiết kế oligonucleotid để tổng hợp gen nhân tạo: Do gen mã hóa Cas9 có kích thước khá lớn (4.107 bp) nên sau khi

tối ưu sẽ chia trình tự gen cần tổng hợp thành 5 vùng gen nhỏ với các điểm cắt của enzyme giới hạn Bsal ở 2 đầu vùng gen. Vector biểu hiện được lựa chọn là pET52b. Trong đó, trình tự gen mã hóa được chèn vào giữa 2 điểm cắt BamHI và NotI. Cas9 tái tổ hợp là một protein dung hợp với đầu N mang đuôi dung hợp Strep II và đầu C mang đuôi dung hợp (His)₁₀. Các mảnh gen thành phần được đưa vào phần mềm trực tuyến DNABworks (<https://hpcwebapps.cit.nih.gov/dnaworks/>) để thiết kế các oligonucleotid thành phần.

+ Tổng hợp và sàng lọc mảnh gen Cas9 bằng PCR: Các mảnh gen thành phần tổng hợp từ các oligonucleotid được thiết kế ở trên bằng kỹ thuật overlap PCR. Sản phẩm PCR được tách dòng bằng cách nối vào vector pJET1.2 và biến nạp vào *E. coli* DH10b. Sau khi sàng lọc bằng kỹ thuật PCR khuẩn lạc, các dòng khuẩn lạc dương tính được nuôi cấy tách chiết plasmid. Việc giải trình tự các plasmid giúp lựa chọn được các dòng plasmid mang trình tự mong muốn.

+ Nối các mảnh gen thành phần vào vector biểu hiện bằng công nghệ Golden gate. Sản phẩm của phản ứng được biến nạp vào *E. coli* BL21(DE3). Sau khi sàng lọc bằng PCR khuẩn lạc, các dòng khuẩn lạc được nuôi cấy, tách chiết plasmid phục vụ việc giải trình tự như trên. Kết quả giải trình tự cho phép lựa chọn được dòng khuẩn lạc mang cấu trúc biểu hiện mong muốn.

- Biểu hiện và tinh sạch protein Cas9 tái tổ hợp:



Hình 1: Sơ đồ thiết kế tối ưu điều kiện biểu hiện và tinh sạch protein Cas9 tái tổ hợp.

* Xử lý số liệu: Bằng các phần mềm Bioedit, Rare codon analysis, Codon optimization.

KẾT QUẢ NGHIÊN CỨU VÀ BÀN LUẬN

1. Kết quả tổng hợp gen mã hóa Cas9 bằng công nghệ Golden gate

* Tối ưu và thiết kế oligonucleotid tổng hợp gen Cas9:

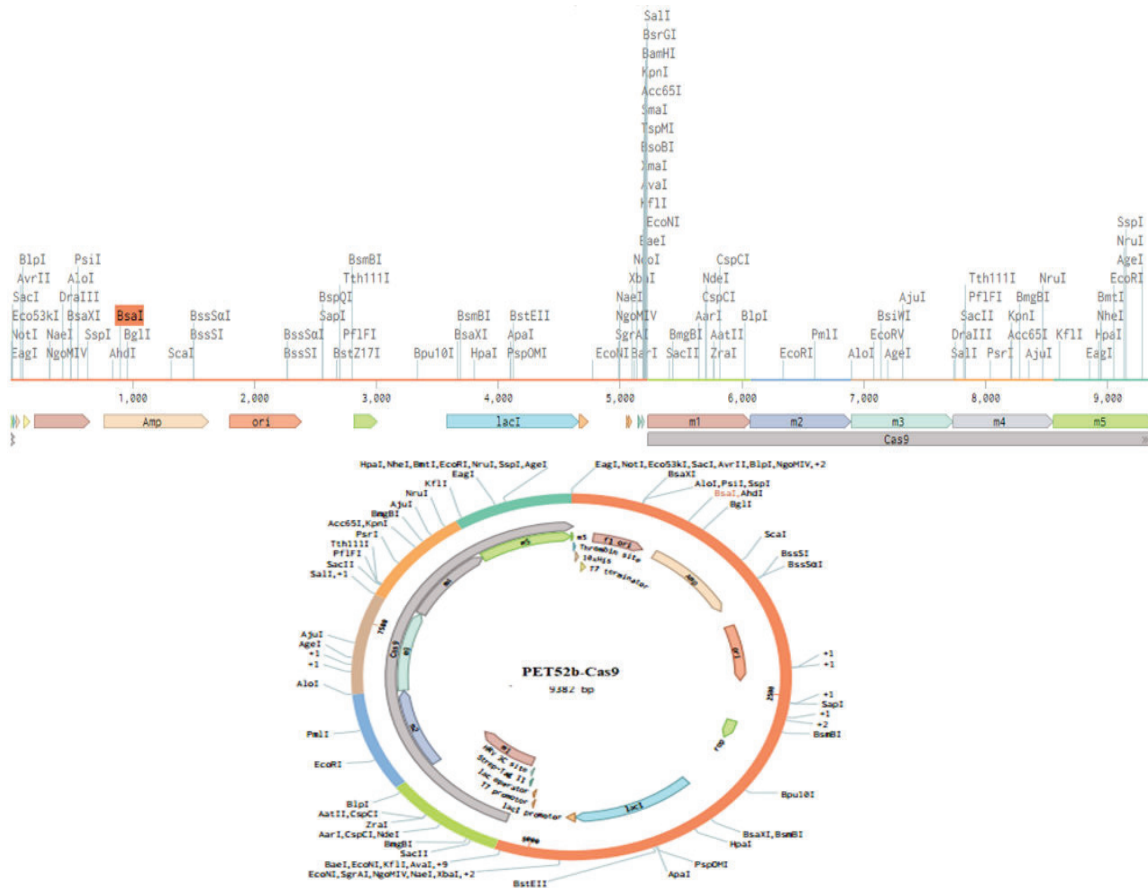
Do gen mã hóa Cas9 có nguồn gốc từ *Streptococcus pyogenes*, khi được biểu hiện trong *E. coli* có thể xảy ra hiện tượng các codon hiếm không được dịch mã tốt trong vật chủ mới. Để kiểm tra, trình tự gen mã hóa cho Cas9 được lấy từ ngân hàng GenBank (mã hiệu AAK33936.1) và phân tích, kiểm tra bằng công cụ

“Genscript Rare Codon Analysis Tools”. Kết quả phân tích cho thấy giá trị CAI của gen này đối với *E. coli* là 0,65. Như vậy, có thể thấy trình tự gen mã hóa Cas9 tự nhiên không thật sự phù hợp để biểu hiện trong *E. coli*. Tiến hành tối ưu hóa trình tự gen mã hóa Cas9 sao cho thích hợp để biểu hiện trong *E. coli* tại địa chỉ <https://eu.idtdna.com/CodonOpt>. Sau khi đã tối ưu hóa codon, trình tự mới thu nhận được phân tích lại bằng công cụ “GenScript Rare Codon Analysis Tool”, kết quả cho thấy giá trị CAI tăng lên 0,85. Đồng thời, tỷ lệ sử dụng các codon hiếm trong gen cũng giảm đáng kể, không còn

các codon có giá trị CFD < 30. Như vậy, trình tự gen sau khi tối ưu phù hợp hơn để biểu hiện trong tế bào *E. coli*. Do gen mã hóa Cas9 có kích thước khá lớn (4.107 bp) nên chia trình tự gen cần tổng hợp thành 5 vùng gen nhỏ với các điểm cắt của enzyme giới hạn BsaI ở 2 đầu vùng gen.

Vector biểu hiện được lựa chọn là pET52b, trong đó trình tự gen mã hóa

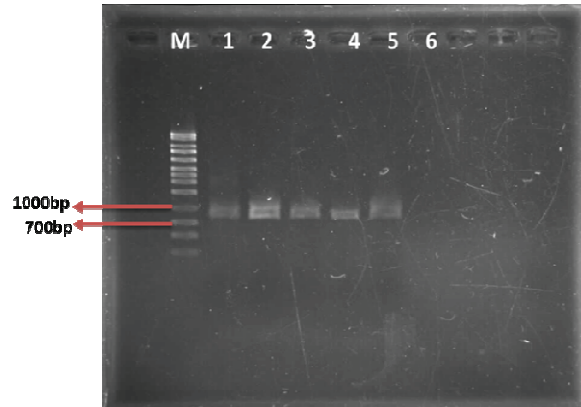
được chèn vào giữa 2 điểm cắt BamHI và NotI. Cas9 tái tổ hợp sẽ là một protein dung hợp với đầu N mang đuôi dung hợp Strep II và đầu C mang đuôi dung hợp (His)₁₀. Điều này giúp đơn giản hóa quá trình tinh sạch Cas9 tái tổ hợp bằng cách sử dụng liên tiếp 2 loại sắc ký ái lực. Thêm trình tự ái nhân dẫn đường ở đầu C giúp protein có khả năng định hướng và hoạt động được trong tế bào động vật.



Hình 2: Trình tự mảnh gen Cas9 lý thuyết sau khi gắn vào vector pET52b.

Sử dụng phần mềm DNAworks thiết kế oligos tổng hợp 5 mảnh gen Cas9, kết quả thu được: Mảnh Cas9.1 có 42 oligos, mảnh Cas9.2 có 40 oligos, mảnh Cas9.3 có 40 oligos, mảnh Cas9.4 có 40 oligos, mảnh Cas9.5 có 42 oligos. Các đoạn oligos được đặt tổng hợp từ hãng Macrogen (Hàn Quốc).

2. Tổng hợp và sàng lọc mảnh gen Cas9 bằng PCR

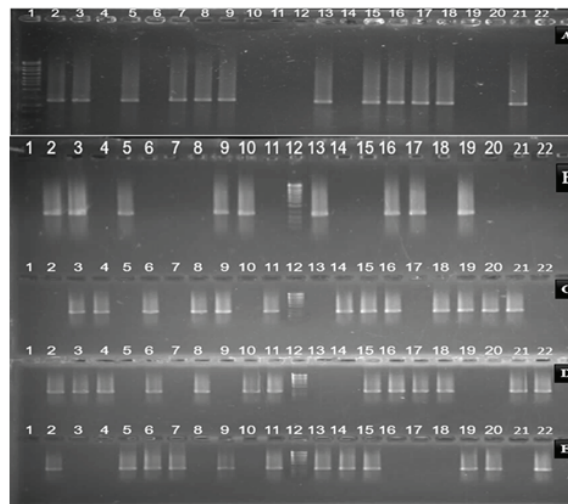


Hình 3: Hình ảnh điện di kết quả khuếch đại từng phân đoạn gen Cas9.

M: HighRanger 1kb DNA Ladder (Norgen);

1: PD1; 2: PD2; 3: PD3; 4: PD4; 5: PD5; 6: Âm tính

Hình ảnh điện di xuất hiện 5 băng vạch kích thước 800 - 900 base, tương ứng kích thước của 5 phân đoạn Cas9 đã thiết kế tổng hợp. Gen Cas9 đã thiết kế tối ưu được tách dòng vào vector PJET 1.2 blunt. Tiến hành biến nạp vào *E. coli* DH10b và sàng lọc bằng kỹ thuật PCR khuẩn lạc.



Hình 4: Kết quả sàng lọc các phân đoạn gen.

A - Kết quả sàng lọc phân đoạn gen Cas9.1; A1: Maker; A22: Âm tính

B - Kết quả sàng lọc phân đoạn gen Cas9.2; B1: Âm tính; B12: Maker

C - Kết quả sàng lọc phân đoạn gen Cas9.3; C1: Âm tính; C12: Maker

D - Kết quả sàng lọc phân đoạn gen Cas9.4; D1: Âm tính; D12: Maker

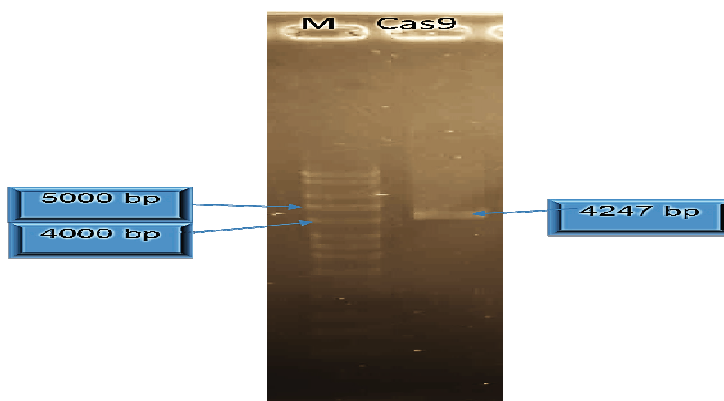
E - Kết quả sàng lọc phân đoạn gen Cas9.5; E1: Âm tính; E12: Maker

Đã sàng lọc được 12 khuẩn lạc mang phân đoạn gen Cas9-1, 9 khuẩn lạc mang phân đoạn gen Cas9-2, 13 khuẩn lạc mang phân đoạn gen Cas9-3, 13 khuẩn lạc mang phân đoạn gen Cas9-4 và 12 khuẩn lạc mang phân đoạn gen Cas9-5.

Lựa chọn 2 dòng vector tái tổ hợp mang mỗi phân đoạn gen Cas9, sau tách chiết được giải trình tự, kết quả: Mảnh 1, 3 và 5 có 2/2 dòng trình tự đúng thiết kế, mảnh 2, 4 có 1/2 dòng trình tự đúng thiết kế (1 dòng thiếu 1 nucleotide ở cuối trình tự).

3. Kết quả nối các mảnh gen thành phần vào vector biểu hiện pET52b bằng công nghệ Golden gate

Các dòng plasmid mang mảnh gen thành phần và pET52b (đã được mở vòng bằng BamHI và SacI) sẽ được trộn với nhau trong phản ứng có chứa đồng thời BsaI và T4 DNA ligase để tổng hợp bằng phương pháp Golden gate. Sản phẩm của phản ứng sau đó được biến nạp vào *E. coli* BL21(DE3). Tiến hành sàng lọc bằng PCR khuẩn lạc, khuếch đại với mỗi T7 promotor và mỗi ngược C9-1R.



Hình 5: Kết quả điện di kiểm tra gen Cas9 được nối từ 5 mảnh.

M: DNA ladder HighRange (Norgen)

Cas9: Sản phẩm PCR khuếch đại gen Cas9 (~4,3kb)

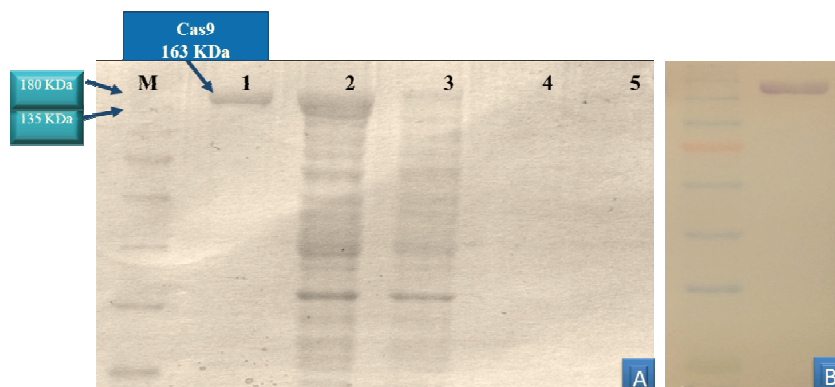
Kết quả điện di thu được 1 băng vạch duy nhất kích thước ~4,3kb. Chứng tỏ gen Cas9 đã được nối thành công vào pET52b và biến nạp thành công vào tế bào vi khuẩn BL21(DE3).

Lựa chọn 5 dòng pET52b-Cas9 để nhân dòng và gửi giải trình tự. Kết quả giải trình tự cho thấy cả 5 dòng đều có trình tự đúng thiết kế ban đầu. Như vậy, qua 4 bước đã tổng hợp thành công gen mã hóa protein Cas9 bằng công nghệ Golden gate mà không cần thu thập nuôi cấy vi khuẩn *Streptococcus pyrogen* để phân lập protein Cas9. Đây là phương pháp mới tạo được gen quan tâm mà không cần sự có mặt của mầm bệnh, góp phần giảm chi phí nghiên cứu cũng như nguy cơ lây nhiễm mầm bệnh.

4. Kết quả biểu hiện và tinh sạch protein rCas 9 trong vi khuẩn *E. coli*

Ở điều kiện nuôi biểu hiện bổ sung 0,2 mM IPTG ở 30°C, qua đêm, vi khuẩn *E. coli* cho hiệu suất sinh tổng hợp protein Cas9 tái tổ hợp cao nhất. Protein Cas9 tái tổ hợp

được xác nhận bằng phương pháp Western blot với kháng thể kháng 10xHis, tinh sạch bằng sắc ký ái lực Ni-NTA và đánh giá độ tinh sạch bằng kỹ thuật điện di SDS-PAGE.



Hình 6: Kết quả tinh sạch protein Cas9 SDS-PAGE (A) và Western blot (B).
M: Ladder Gangnam Stain; 1: Dịch rửa giải sau khi tinh sạch;
2: Dịch trước khi tinh sạch; 3: Dịch sau khi qua cột; 4, 5: Dịch rửa

Kết quả điện di SDS-PAGE và Western blot cho thấy đã biểu hiện thành công cũng như tinh sạch protein Cas9 tái tổ hợp với kích thước khoảng 163 KDa (do gắn thêm đuôi Histag). Lượng protein rCas9 thu được có nồng độ 4,1 mg/μl và độ tinh sạch $\geq 98\%$.

KẾT LUẬN

Chúng tôi đã cấu trúc thành công vector tái tổ hợp mang gen Cas9 (pET52b-Cas9) mã hóa cho protein Cas9 có nguồn gốc từ *Streptococcus pyrogenes*; tạo thành công dòng tế bào *E. coli* BL21(DE3) mang vector pET52b-Cas 9 có khả năng biểu hiện protein Cas9 tái tổ hợp ở pha tan và thu nhận được Cas9 tái tổ hợp với độ tinh sạch $\geq 98\%$. Đã tạo được nguồn cung cấp Cas9 tái tổ hợp phục vụ cho các nghiên cứu tiếp theo nhằm đánh giá hoạt tính endonuclease của protein Cas9, hướng tới việc ứng dụng phức hợp CRISPR/Cas9 trong y sinh học.

TÀI LIỆU THAM KHẢO

1. Jackson Campbell, Rece, Urry, Cain, Wasserman, Minorsky. Chapter 20: Biotechnology. BIOLOGY (8th edition) 2008: 396-425.

2. Chen B, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell 2013; 155:1479-1491.

3. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. Science 2013; 339(6121):819-823.

4. George M Church, et al. CRISPR-Cas9 system: Opportunity and concern. doi:10.1373/clinchem.2016.263186.

5. Gilbert LA, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 2013; 154:442-451.

6. Harrison MM, et al. A CRISPR view of development. Genes Dev 2014; 28:1859-1872.

7. Ran FA, et al. Genome engineering using the CRISPR-Cas9 system. doi:10.1038/nprot.2013.143.